4/PRTS

09/743430

JC07 Rec'd PCT/PTO 0 8 JAN 2001

Description

Method and arrangement for the determination of a measure of similarity between a first structure and at least one predetermined second structure

The invention relates to the computer-aided determination of a measure of similarity between a first structure and at least one predetermined second structure.

Such a method is disclosed in [1]. In this image processing method, in the context of a motion estimation between two chronologically successive images for an image block to be coded, in the chronologically preceding image, an image area is sought which is as similar as possible to the image block to be coded. This is done by determining a sum of the square differences of the pixel-assigned coding information (brightness value or color value) of the pixels of the image block to be coded and of the pixels in an investigated area in the chronologically preceding image. This means that for the comparison of two structures in this approach from [1], the entire first structure (image block to be coded) is compared with a second structure (area in the chronologically preceding image) in its entirety.

This procedure is extremely costly with regard to the memory space required for storing the structures, and also with regard to the computing time required for determining the measure of similarity (in this case, the measure of similarity is the sum of the square differences).

The following procedure is disclosed in [2]: for the comparison of a structure, the elements of the structure are interpreted as polygonal progressions and

the polygonal progressions are compared in their entirety, angles between interconnected elements of the polygonal progressions being taken into account.

This approach manifests the disadvantage, in particular, that it is not robust relative to recordings made by a robot for its orientation in a predetermined space. The space is in this case represented by a predetermined, stored map.

If some elements of the polygonal progression are not recorded by the robot, then the method disclosed in [2] is not robust enough to lead to results of sufficient quality. Moreover, the [lacuna] for the comparison of the recorded image with the map in order to compare structures with one another is extremely computationally intensive.

Consequently, the invention is based on the problem of determining a measure of similarity between a first structure and at least one predetermined second structure which is robust with respect to possible recording errors and can be carried out more rapidly, with less computation time being required, than the known methods.

The problem is solved by means of the method in accordance with patent claim 1 and also by means of the arrangement in accordance with patent claim 27.

In the case of the method, in each case at least one base element is defined for the first structure and the second structure. Surroundings-related information is assigned to each of the base elements, which surroundings-related information characterizes the corresponding base element. The measure of similarity, which describes the similarity between the first structure and the second structure, is determined for

the first structure and the second structure. The measure of similarity is determined in a manner dependent on the base elements and on the surroundings-related information assigned to the base elements.

The arrangement has a processor which is set up in such a way that the following steps can be carried out:

- in each case at least one base element is defined for the first structure and the second structure,
- surroundings-related information is assigned to each of the base elements,
- the surroundings-related information characterizes the corresponding base element,
- the measure of similarity, which describes the similarity between the first structure and the second structure, is determined for the first structure and the second structure, and
- the measure of similarity is determined in a manner dependent on the base elements and on the surroundings-related information assigned to the base elements.

The invention is distinguished by the fact that, compared with the prior art, considerably less computation time is required for carrying out the determination of the measure of similarity, and also by increased robustness with respect to possible sensor errors or modeling errors.

Preferred developments of the invention emerge from the dependent claims.

The structures may be contained in a map which is recorded as a scene from the surroundings using a recording means, for example a laser scanner or a camera. In this case, it is necessary to compare the recorded structures with a stored map in order thus to

perform orientation or to construct a map which can be used for orientation.

The accuracy of the method is increased if the measure of similarity is determined on the basis of a plurality of base elements and the respectively assigned surroundings-related information thereof.

A further refinement of the invention provides for at least a portion of the base elements to be a line, a section or else a point.

The surroundings-related information may be formed by further base elements and the geometrical arrangement thereof relative to the base elements.

The base elements may have different forms.

A development of the invention whereby at least a portion of the surroundings-related information is formed in such a way that it is invariant with respect to errors which occur when constructing the map has the effect of achieving a further qualitative improvement in the results.

A further simplification of the invention can be achieved in that, in a preferred development, the further base elements are grouped into a plurality of surroundings-related information types containing surroundings-related information features which are each assigned to a surroundings-related information type.

The surroundings-related information features assigned to a surroundings-related information type may be stored having been sorted in a predeterminable manner in a list.

The sorting can be carried out in such a way that it is invariant with respect to the errors - explained above - when constructing the map.

The measure of similarity is preferably determined by means of dynamic programming.

The structures may each describe a physical object, for example a recorded space, or else a recorded scene, and also, by way of example walls, doors, or any other objects.

In an alternative, it is likewise possible for the first structure to describe a physical object and the second structure to describe a model of a physical object.

Furthermore, the structures may represent data structures in a database.

The invention can preferably be used for the orientation of a mobile autonomous apparatus or else for the determination of a map for the orientation of said apparatus.

In a preferred development, the autonomous apparatus is a robot.

Exemplary embodiments of the invention are illustrated in the figures and are explained in more detail below.

In the figures,

Figures 1a and 1b show a sketch of a passageway in which a robot is intended to orient itself (figure 1a), and also a symbolic sketch of the recordings of the robot and its conversion into a map, an error in the determination of the map

and its effects on the mapping of the passageway relative to the actual passageway from figure 1a being illustrated (figure 1b);

Figure 2 shows a sketch of a robot with recording means;

Figures 3a to 3c show sketches in each case of a base element with different surroundings-related information types and surroundings-related information features;

Figure 4 shows a sketch illustrating an application of the method, in which a structure represents a model of a physical object;

Figure 5 shows a flow diagram illustrating method steps of an exemplary embodiment.

Figure 2 shows a robot 201 with a plurality of laser scanners 202. The laser scanners 202 record images of the surroundings of the robot 201 and feed the images to a computing unit 203 via connections 204, 205.

The image signals are fed to a memory 208 via an input/output interface 206, which is connected via a bus 207 to the memory 208 and also to a processor 209.

The method described below is carried out in the processor 209. Consequently, the processor 209 is set up in such a way that the method steps described below can be carried out.

Figure 1a symbolically shows a map 101 representing a passageway 102. The robot 201 moves through the passageway and records images of its surroundings using the laser scanners 202. It records walls 103 in the process. The robot 201 records images of its

surroundings at different times, an image of the entire map 101 thereby being produced.

In the passageway 102 there are, moreover, obstacles 104 in the form of shelves or else cabinets or similar items which project into the passageway 102.

Corners 105, 106, 107 of the passageway 102 are interpreted as the starting point and end point of a wall, which is stored in the form of a section segment.

Figure 1b illustrates the map from **figure 1a**, not when ideal recordings are made, as is assumed in the case of the situation illustrated in **figure 1a**, but when errors happen in the course of the recording by the robot 201.

The robot 201 moves in the passageway 102 and records images of its surroundings at periodic intervals. The robot 201 orients itself on the basis of the recorded images and also of the stored map 101.

The orientation takes place in such a way that the robot 201 feeds the images to the processor 209. In the processor 209, a similarity comparison of elements of the recorded image with elements of the stored, predetermined map 101 is determined and an attempt is made to determine from this the present position of the robot 201.

The robot 201 is situated at a position 110 and records an image area 111 using its laser scanner. It attempts to match said image area 111 to the stored map 101 in order thus to determine information for its orientation.

This corresponds to the comparison of a first structure, which characterizes the recorded image area

111, with at least one predetermined second structure from the predetermined, stored map 101.

The following method, which is illustrated in **figure 5**, is carried out for this purpose.

In a first step 501, base elements are extracted from the recorded image 111 by the processor 209.

A base element is to be understood as a section having a starting point and an end point, which in each case represents a wall in the passageway 102. Further base elements are points or lines of predeterminable form.

The extraction is effected using known image processing methods.

After the extraction of the base elements, the image is present, symbolically represented by a set of defined base elements.

Surroundings-related information is assigned to each base element. The surroundings-related information characterizes the corresponding base element and enables the identification of the respective base element within a set of all the base elements.

Such a base element 301 with surroundings-related information 302 assigned to the base element is illustrated in each case in **figure 3a** to **figure 3c**.

The surroundings-related information is formed by a set of further base elements and the geometrical arrangement thereof relative to one another and also to the base element 301 itself.

The surroundings-related information assigned to the base element 301 is formed in such a way that it is as

far as possible invariant with respect to errors which can occur when the map 101 is constructed by the robot 201.

It has been found that a pair of orthogonal base elements in the form of sections are particularly well suited to the method. In this case, it should be noted that exact orthogonality of the base elements is not important, rather a tolerance can readily be accepted.

The surroundings-related information assigned to the base element 301 is the distance between the points of intersection of the parallel base elements with the base element 301, designated by Dx in **figure 3a**.

Furthermore, a first angle W1, which denotes an angle of intersection of a first further base element 303, having a length L1, with the base element 301, is stored as surroundings-related information.

Furthermore, a second angle W2, which denotes an angle of intersection of the second further base element 304 with the base element 301, and also the length L2 of the second further base element 303 are assigned to the base element 301 as surroundings-related information.

Further surroundings-related information assigned to the base element 301 is an indication of the starting point and/or of an end point and thus also an orientation in each case of the first and/or second base element 303, 304.

The surroundings-related information is stored as a list assigned to the base element 301. The list is sorted in a predeterminable manner.

The pair of orthogonal base elements, as illustrated in **figure 3a** as surroundings-related information, forms a surroundings-related information type.

The above-described individual elements assigned to the base element 301 as surroundings-related information form surroundings-related information features which are each assigned to the surroundings-related information type.

A second surroundings-related information type is a further base element 310 parallel to the base element 301 (cf. **figure 3b**).

Once again, exact parallelism of the further base element 310 with respect to the base element 301 is not necessary. A distance Dy between the base element 301 and the further, parallel base element 303 and also a third angle W3, between the exact parallel line 311 with respect to the first base element 301, shifted by the distance Dy, and the actual position of the further, parallel base element 310 are stored as surroundings-related information features.

Figure 3c shows a further surroundings-related information type in the form of points 320, 321, which denote points of a line structure 322 which lie the nearest to the base element 301. In this case, a distance between said points 320, 321 (designated as Dz) and also the shortest distances N1, N2 of the points 320, 321 from the base element 301 are stored as surroundings-related information features.

In the predetermined, stored map 101, surroundings-related information items are in each case assigned to the base elements in the same way. Consequently, the stored map 101 has a set of base elements with, in each case assigned to the base elements, surroundings-

related information in the form of surroundings-related information types with surroundings-related information features assigned to the surroundings-related information types.

Thus, in a second step 502, the surroundings-related information items are in each case assigned to the base elements contained in the image area 111 and also to the base elements contained in the map 101.

For each base element 301, in a further step 503, a value of a measure of similarity is formed with all the further base elements.

The measure of similarity is explained in more detail below.

In this exemplary embodiment, it is assumed that a total value U of the surroundings-related information assigned in each case to the base element 301 is produced according to the following specification:

$$U = (OP, P, MP),$$

where

- OP designates the surroundings-related information features which are formed by pairs of further base elements oriented perpendicularly to one another,
- P designates the surroundings-related information features, formed by parallel base elements, and
- MP designates the surroundings-related information features of the punctiform surroundings-related information types.

The surroundings-related information features are present in the form of sorted lists.

Let

$$v : U \times U \rightarrow \Re_0^+$$

be a formal definition of a comparison function.

Using the comparison function v, a comparison value is calculated for a pair of surroundings-related information items assigned in each case to two base elements. The higher the comparison value, the better the two surroundings-related information features of the base elements correspond to one another. For the definition of the comparison function v, the following three functions vOP, vP, vMP are defined:

$$vOP : OP \times OP \rightarrow \Re_0^+$$

$$vP : P \times P \rightarrow \Re_0^+$$

$$vMP : MP \times MP \rightarrow \Re_0^+,$$

where vOP describes a comparison value for surroundings-related information features of the surroundings-related information type with perpendicular further base elements and, analogously vP describes a comparison value of surroundings-related information features of the surroundings-related information type with parallel base elements. vMP describes a comparison value which determines surroundings-related information features of the surroundings-related information type with points as surroundings-related information features.

The comparison function v is defined as the weighted sum of the functions vOP, vP and vMP according to the following specification.

v(U1, U2) = aOP * vOP(OP1, OP2) + aP * vP(P1, P2) + aMP * * vMP(MP1, MP2).

The values aOP, aP and aMP in the numerical interval [0,1] are designated as weight values.

With the weight values aOP, aP and aMP, account is taken of the different significances of the individual surroundings-related information types with regard to the measure of similarity. It has been found that the surroundings-related information type of the pairs of orthogonal further base elements OP has a greater meaningfulness with regard to the measure of similarity than the surroundings-related information type of the parallel further base elements P and the latter in turn has a greater meaningfulness than the surroundings-related information type with points as surroundings-related information features.

For each function vOP, vP, vMP a method of dynamic programming is in each case carried out for each base element and the surroundings-related information features thereof, as a result of which an intermediate similarity value is formed.

This is done in each case for each function vOP, vP, vMP using the following cost function $D_{i,j}$:

$$D_{i,j} = \min \begin{Bmatrix} D_{i-1,j} + \delta \\ D_{i-1,j-1} + \mu \\ D_{i,j-1} + \delta \end{Bmatrix}$$

where

- δ, a pre-determinable cost value which occurs if a surroundings-related information feature of the recorded image area cannot be assigned to a

surroundings-related information feature of the stored map 101,

- $\mu = \dfrac{\delta}{\lambda}$, where

$$-\lambda = \prod_{k=1}^{n}\left(\max\left\{0, 1 - \frac{\left|a_{k,i} - a_{k,j}\right|}{MaxErr_k}\right\}\right).$$

In this case

- k designates an index which unambiguously designates each surroundings-related information type which is taken into account in the context of the dynamic programming,

- n designates the number of base elements taken into account,

- $a_{k,i}$ and $a_{k,j}$ designate the individual surroundings-related information features which are stored in the sorted list of the respective surroundings-related information types, $a_{k,i}$ designating a surroundings-related information feature of a base element of the image area 111 and $a_{k,j}$ designating a surroundings-related information feature of a base element of the map 101,

- $MaxErr_k$ designates a predeterminable value specific to each surroundings-related information type.

The cost value δ should be determined empirically in such a way that, for the given application,

$2 \cdot \delta > \mu$ if the assignment is correct, and
$\mu > 2 \cdot \delta$ if the assignment is not correct.

The following ratio of the individual weight values has been found to be advantageous:

aOP : aP : aMP = 3 : 2 : 1.

The result of the comparison fucntion v forms a value of the measure of similarity which describes the similarity between the first structure in the image area 111 and the second structure in the map 101 (step 503).

In a further step 504, the pair of base elements from the first structure and/or the second structure is selected, which pair has the highest value of the intermediate similarity value and is thus the most similar to one another.

For the selected base elements, a canonical coordinate system is formed in the respective map, the abscissa of which system is formed by the respective base element (step 505).

In a further step 506 a mapping measure is subsequently determined. The mapping measure is used to determine, for the selected base elements, what magnitude of a translation or rotation is necessary in order to map the coordinate system for the base element of the first structure in each case onto a coordinate system of a base element of a further structure.

Thus, step 506 clearly determines in each case the extent to which the coordinate system of the selected base element of the first structure must be shifted or "rotated" in order to "match" the coordinate system of the selected base element of a respective further structure.

The area selected in the predetermined map is that area whose mapping measure and/or whose measure of similarity is minimal compared with the coordinate system for the base element of the first structure.

Proceeding from the selected base element, further base elements are selected in pairs (i.e. in each case a base element of the first structure and a base element of the second structure), whose values of the measures of similarity are greater than a predeterminable threshold value.

The robot 201 now knows where it is situated within the map 101.

Consequently, in a final step 507, that area in the predetermined map 101 in which the robot 201 is situated is determined.

A number of alternatives to the exemplary embodiment described above are presented below:

The method described above can also be used for the general comparison of two structures, for example for the comparison of a recorded image with a model of a physical object.

One example of this is to be seen in the model of a door handle, which door handle is to be gripped by the robot 201.

A model 401 of a door handle 402 is stored in the memory 208 of the computing unit 203 of the robot 201.

The structure comparison is effected in the manner described above, the scanner of the robot 201 recording images of its surroundings and searching for a structure which is similar to the structure of the

door-handle model 401. If such a structure is determined, then a gripping arm 403 of the robot 201 can grip the door handle 401, which is fitted to a door 404.

A further possibility for using the method is to be seen in the field of databases. Databases likewise have a structure, in which the data are stored. Consequently, the structure of the stored database can clearly be compared with a sought structure in the manner described above, and, as a result of this geometrical interpretation of the structure of a database, it is possible to determine a segment of the database in order to ensure a high reliability of the search results in the database in the context of a search.

Furthermore, the method can also be used for the progressive construction of the map 101 by the robot 201. This method then serves for checking when the robot 201 reaches a location where it has already been. In this case, the robot in each case compares the recording of an image with the stored structure data of a map 101 that is being constructed.

The method can also generally be used in the context of pattern recognition or image processing, the computation time required for carrying out the method being considerably reduced compared with the known methods in pattern recognition.

The following publications were cited in the context of this document:

[1] M. Bierling, Displacement Estimation by Hierarchical Blockmatching, SPIE, Vol. 1001, Visual Communications and Image Processing '88, pp. 942 - 951, 1988

[2] O. Karch and H. Noltemeier, Autonome Mobile Systeme [Autonomous Mobile Systems] 1996, G. Schmidt and F. Freyberger, (Eds.), Zum Lokalisationsproblem für Roboter [Regarding the Localization Problem for Robots], Springer Verlag, ISBN 3-54061-751-5, pp. 128 - 137, 1996